

Building a Culture of Data Management within USGS Chesapeake Bay Studies

Cassandra Ladino (ccladino@usgs.gov), Eastern Geographic Science Center, Reston, VA

People

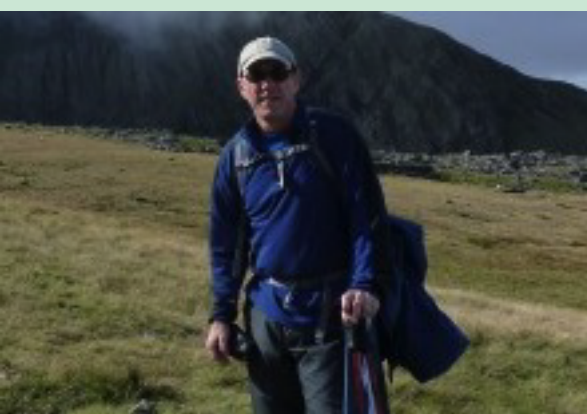
“It takes a village to raise a child”...

Data Management can’t be the effort of one, it takes a community to build a culture of practice. Here are some of the people creating a culture of Data Management within USGS Chesapeake Studies.



Fish Health

Vicki Blazer, Research Fish Biologist, is compiling a multitude of fish sample data sets that date back to 2004. Vicki is also thinking about a sustainable method for compiling this data going into the future.



Contaminants

Patrick Phillips, Supervisory Hydrologist, is developing a method to pull subsets of water and sediment samples tested for pesticides, hormones, wastewater, and pharmaceuticals from National Water Information System (NWIS) to combine with the Fish Health information. Pat is also rescuing data from 2007/ 2008 that has not been entered into NWIS.

David Alvarez, Supervisory Research Chemist, is working on compiling passive sampler data from 2005 to present to be combined with the Fish Health information.



Water Quality Trends

Joel Blomquist, Hydrologist, wants to improve internal access and sharing of the information that his water quality trends group produces annually.

Solutions

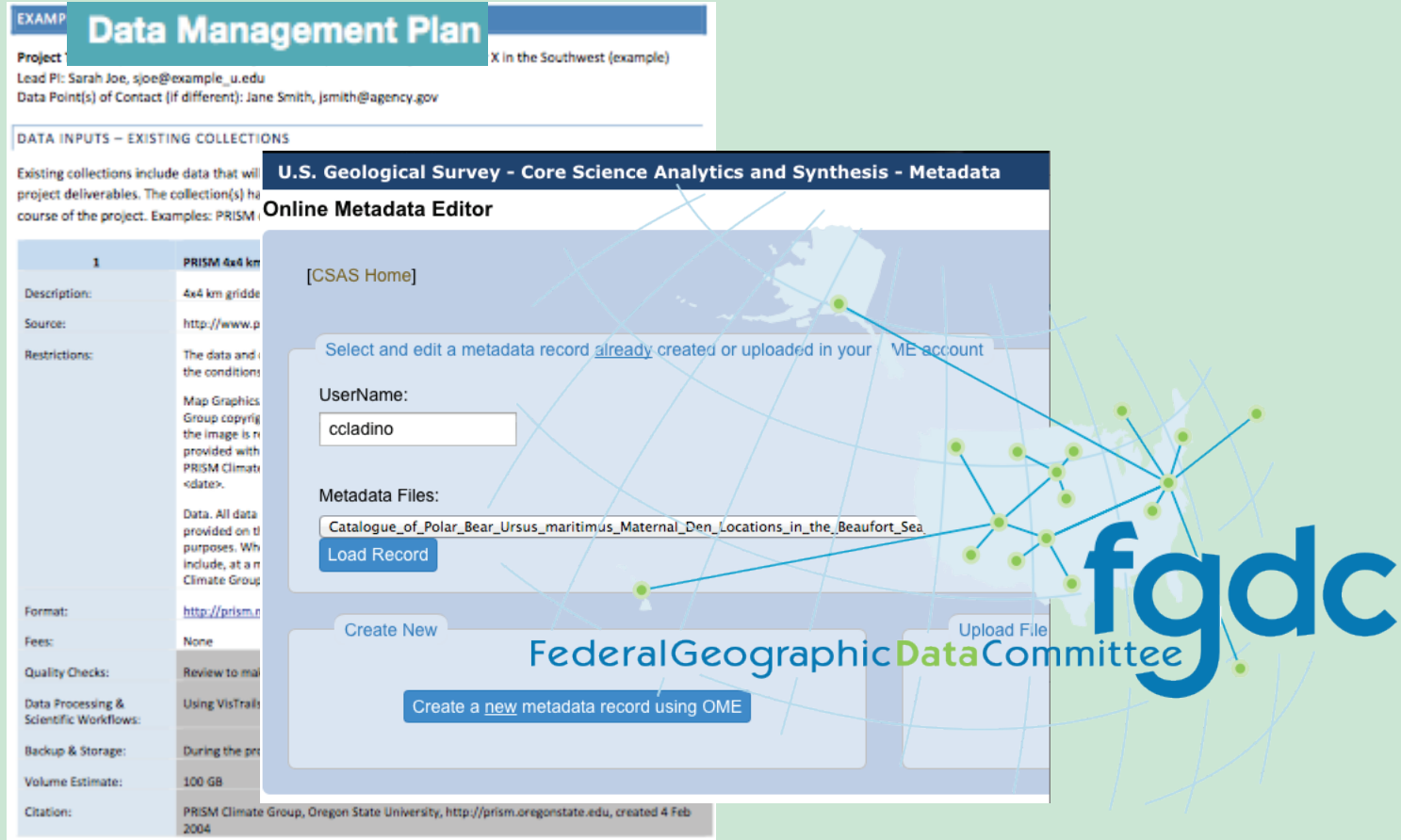
“Where there’s a will, there’s a way”...

Technology sololutions are helping to reduce the burden of changing requirements. To the right are some of our current solutions to data curation, producing documentation (Data Management Plans and Metadata), and building integrated databases that benefit scientists directly.



Hire a Data Manager

The Environmental Health Mission Area’s Endocrine Disrupting Chemicals group recognized the importance of building an integrated science theme database and meeting new Fundamental Science Practices. The key to meeting those goals is to implement Data Management best practices. Scientist do not have budgets to cover these tasks, nor do information technology team, as a result a data manageris being hired and trained based on resources highlighted through the Community for Data Integration and the USGS Data Management website.



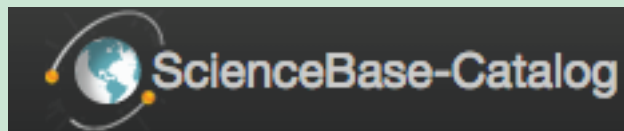
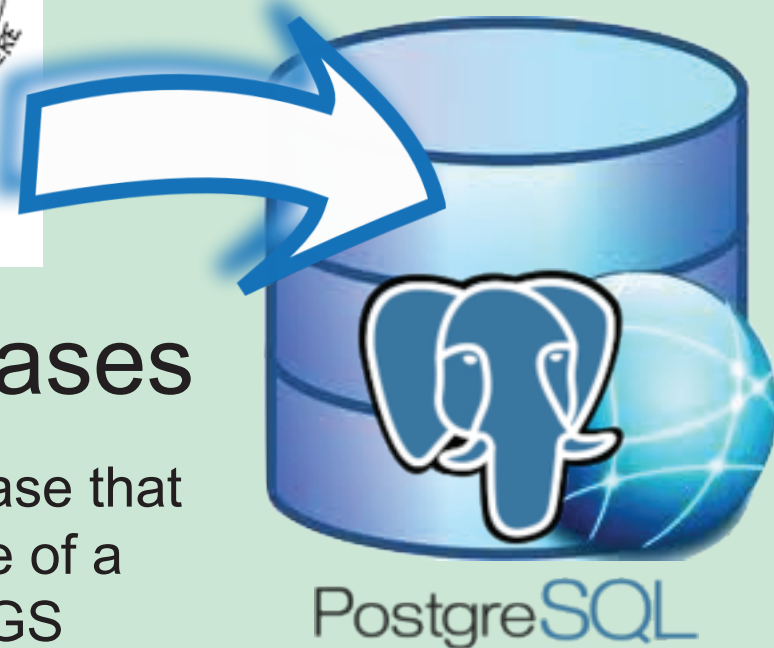
Implement Standards

The Data Manager will work with the scientists this summer to develop Data Management Plans based on USGS NCCWSC standards and create metadata using the USGS Online Metadata Editor. This will help keep project workflows and data sets consistent year to year.



Develop Databases

EGSC is building a database that will serve as the backbone of a data infrastructure for USGS Chesapeake Bay studies. This year the focus is to integrate Fish Health and Water Contaminants information.



Utilize ScienceBase

Scientists are encouraged to use the ScienceBase USGS Chesapeake Bay community to make published datasets available via a persistent location and reduce data redundancy among projects.

```
library(RPostgreSQL)
drv <- dbDriver("PostgreSQL")
con <- dbConnect(drv, user="user",password="psswr",
dbname="mydb", host="remotehost",port="3333")
```



Provide Data Services

The EGSC Web Development team created a web accessible database and is providing code snippets to encourage use of remote data sets vs. local copies.

Discovery

“The Times They Are a-Changin”...

Changing project work flows and old “normals” has encouraged innovation within the Chesapeake community. Here are some of the innovations that we’ve discovered are improving our ability to adopt Data Management best practices.



Creating partnerships and communities of practice is a promising way to promote culture change and reduce the burden of adopting Data Management best practices. Partnering with the Environmental Health Mission Area’s Endocrine Disrupting Chemicals program and Aquatic GAP has enabled us to create more solutions to solve challenges.



Good communication between technical teams and scientists leads to learning on both sides. The contaminants team members have learned a lot about data curation by understanding the requirements of database development. The technical team developing the database has learned a lot about the special science considerations of each data set to ensure the database maintains data integrity.



When Data Management and its benefits are understood, scientists see new possibilities for the data they produce. The water quality trends group is excited about using new databases to more efficiently analyze and publish their annual data series.